Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L6	3	(synthetic near2 (hyperlink or hyper?link or url))	US-PGPUB; USPAT; EPO; JPO; IBM_TDB	OR	ON	2005/06/21 10:34
L7	180	((synthetic or artificial or generat$4 or construct$3 or build$3 or compos$3) near2 (hyperlink or hyper?link or url)) and (crawl$3 or spider$3)	US-PGPUB; USPAT; EPO; JPO; IBM_TDB	OR	ON	2005/06/21 11:19
L8	53	((synthetic or artificial or generat$4 or construct$3 or build$3 or compos$3) near2 (hyperlink or hyper?link or url)) same (crawl$3 or spider$3)	US-PGPUB; USPAT; EPO; JPO; IBM_TDB	OR	ON	2005/06/21 10:36
L9	30	L8 and dynamic$5	US-PGPUB; USPAT; EPO; JPO; IBM_TDB	OR	ON	2005/06/21 10:37
L10	37	L7 and (dynamic near2 (page or content))	US-PGPUB; USPAT; EPO; JPO; IBM_TDB	OR	ON	2005/06/21 11:19
L11	11	L10 and @ad<="20001218"	US-PGPUB; USPAT; EPO; JPO; IBM_TDB	OR	ON	2005/06/21 12:06
L12	123	(dynamic near1 (url or link or hyperlink or hyper?link or page)) same (index$3 or crawl$3 or spider$3)	US-PGPUB; USPAT; EPO; JPO; IBM_TDB	OR	ON	2005/06/21 11:58
L13	43	L12 and @ad<="20001218"	US-PGPUB; USPAT; EPO; JPO; IBM_TDB	OR	ON	2005/06/21 12:06

Published before December 2000
Terms used **dynamic page crawl url**

Found **52** of **111,743**

Sort results by: relevance
Display results: expanded form
Save results to a Binder
Search Tips
Open results in a new window

Try an Advanced Search
Try this search in The ACM Guide

Results 1 - 20 of 52 Result page: **1** 2 3 next

Relevance scale

1 Crawler-Friendly Web Servers
Onn Brandman, Junghoo Cho, Hector Garcia-Molina, Narayanan Shivakumar
September 2000 **ACM SIGMETRICS Performance Evaluation Review**, Volume 28 Issue 2
Full text available: pdf(513.04 KB) Additional Information: full citation, abstract, index terms

In this paper we study how to make web servers (e.g., Apache) more crawler friendly. Current web servers offer the same interface to crawlers and regular web surfers, even though crawlers and surfers have very different performance requirements. We evaluate simple and easy-to-incorporate modifications to web servers so that there are significant bandwidth savings. Specifically, we propose that web servers export meta-data archives decribing their content.

2 Indexing and retrieval of scientific literature
Steve Lawrence, Kurt Bollacker, C. Lee Giles
November 1999 **Proceedings of the eighth international conference on Information and knowledge management**
Full text available: pdf(985.22 KB) Additional Information: full citation, abstract, references, citings, index terms

The web has greatly improved access to scientific literature. However, scientific articles on the web are largely disorganized, with research articles being spread across archive sites, institution sites, journal sites, and researcher homepages. No index covers all of the available literature, and the major web search engines typically do not index the content of Postscript/PDF documents at all. This paper discusses the creation of digital libraries of scientific literature on the web, incl ...

3 Mining multimedia data
Osmar R. Zaïane, Jiawei Han, Ze-Nian Li, Jean Hou
November 1998 **Proceedings of the 1998 conference of the Centre for Advanced Studies on Collaborative research**
Full text available: pdf(377.84 KB) Additional Information: full citation, abstract, references, citings, index terms

Data Mining is a young but flourishing field. Many algorithms and applications exist to mine different types of data and extract different types of knowledge. Mining multimedia data is, however, at an experimental stage.We have implemented a prototype for mining high-level multimedia information and knowledge from large multimedia databases. MultiMedia Miner has been designed based on our years of experience in the research and development of a

relational data mining system, DBMiner, in the Inte ...

Keywords: data cube, data mining, data warehousing, image analysis, information retrieval, multimedia, world-wide web

4 Topical locality in the Web
Brian D. Davison
July 2000 **Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval**

Full text available: pdf(771.77 KB) Additional Information: full citation, abstract, references, citings, index terms

Most web pages are linked to others with related content. This idea, combined with another that says that *text in, and possibly around, HTML anchors describe the pages to which they point*, is the foundation for a usable World-Wide Web. In this paper, we examine to what extent these ideas hold by empirically testing whether topical locality mirrors spatial locality of pages on the Web. In particular, we find that the likelihood of linked pages having similar textual content to be ...

5 Information retrieval on the web
Mei Kobayashi, Koichi Takeda
June 2000 **ACM Computing Surveys (CSUR)**, Volume 32 Issue 2

Full text available: pdf(213.89 KB) Additional Information: full citation, abstract, references, citings, index terms

In this paper we review studies of the growth of the Internet and technologies that are useful for information search and retrieval on the Web. We present data on the Internet from several different sources, e.g., current as well as projected number of users, hosts, and Web sites. Although numerical figures vary, overall trends cited by the sources are consistent and point to exponential growth in the past and in the coming decade. Hence it is not surprising that about 85% of Internet user ...

Keywords: Internet, World Wide Web, clustering, indexing, information retrieval, knowledge management, search engine

6 Organizing topic-specific web information
Sougata Mukherjea
May 2000 **Proceedings of the eleventh ACM on Hypertext and hypermedia**

Full text available: pdf(183.02 KB) Additional Information: full citation, references, citings, index terms

Keywords: World-Wide Web, abstraction hierarchy, graph algorithms, information visualization, topic management

7 Database techniques for the World-Wide Web: a survey
Daniela Florescu, Alon Levy, Alberto Mendelzon
September 1998 **ACM SIGMOD Record**, Volume 27 Issue 3

Full text available: pdf(1.79 MB) Additional Information: full citation, citings, index terms

8 Toward a Dexter-based model for open hypermedia: unifying embedded references

and link objects
Kaj Grønbæk, Randall H. Trigg
March 1996 **Proceedings of the the seventh ACM conference on Hypertext**
Full text available: pdf(1.31 MB) Additional Information: full citation, references, citings, index terms

Keywords: Dexter hypertext reference model, dynamic hypermedia, embedded links, generic links, link objects, open hypermedia

9 Defining logical domains in a web site
Wen-Syan Li, Okan Kolak, Quoc Vu, Hajime Takano
May 2000 **Proceedings of the eleventh ACM on Hypertext and hypermedia**
Full text available: pdf(152.26 KB) Additional Information: full citation, references, citings, index terms

Keywords: WWW, domain boundary, link structures, logical domain, site map

10 Performance limitations of the Java core libraries
Allan Heydon, Marc Najork
June 1999 **Proceedings of the ACM 1999 conference on Java Grande**
Full text available: pdf(873.12 KB) Additional Information: full citation, references, citings, index terms

Keywords: Java class libraries, Java performance, web crawling

11 Constructing, organizing, and visualizing collections of topically related Web resources
Loren Terveen, Will Hill, Brian Amento
March 1999 **ACM Transactions on Computer-Human Interaction (TOCHI)**, Volume 6 Issue 1
Full text available: pdf(303.62 KB) Additional Information: full citation, abstract, references, citings, index terms

For many purposes, the Web page is too small a unit of interaction and analysis. Web sites are structured multimedia documents consisting of many pages, and users often are interested in obtaining and evaluating entire collections of topically related sites. Once such a collection is obtained, users face the challenge of exploring, comprehending and organizing the items. We report four innovations that address these user needs: (1) we replaced the Web page with the Web site
Keywords: cocitation analysis, collaborative filtering, computer supported cooperative work, information visualization, social filtering, social network analysis

12 WebCQ-detecting and delivering information changes on the web
Ling Liu, Calton Pu, Wei Tang
November 2000 **Proceedings of the ninth international conference on Information and knowledge management**
Full text available: pdf(835.31 KB) Additional Information: full citation, references, citings, index terms

13 Finding replicated Web collections

Junghoo Cho, Narayanan Shivakumar, Hector Garcia-Molina
May 2000 **ACM SIGMOD Record , Proceedings of the 2000 ACM SIGMOD international conference on Management of data**, Volume 29 Issue 2

Full text available: pdf(332.72 KB) Additional Information: full citation, abstract, references, citings, index terms

Many web documents (such as JAVA FAQs) are being replicated on the Internet. Often entire document collections (such as hyperlinked Linux manuals) are being replicated many times. In this paper, we make the case for identifying replicated documents and collections to improve web crawlers, archivers, and ranking functions used in search engines. The paper describes how to efficiently identify replicated documents and hyperlinked document collections. The challenge is to identify these replicas ...

14 Navigating in information spaces: Rapid-fire image previews for information navigation
Kent Wittenburg, Wissam Ali-Ahmad, Daniel LaLiberte, Tom Lanning
May 1998 **Proceedings of the working conference on Advanced visual interfaces**

Full text available: pdf(3.94 MB) Additional Information: full citation, abstract, references, citings

In this paper we consider the role of rapid-fire presentation of images in the service of navigation in information spaces. We presume a model of information navigation in which the user performs a cycle of (pre)viewing, selecting, and moving. Our hypothesis is that images presented to the user in rapid succession can significantly enhance the previewing step, thus optimizing the selection step and improving navigability. We discuss two prototypes for navigation tools in Web information spaces i ...

Keywords: images, information navigation, previewing, visualization

15 Ready for prime time: pre-generation of web pages in TIScover
Birgit Pröll, Heinrich Starck, Werner Retschitzegger, Harald Sighart
November 1999 **Proceedings of the eighth international conference on Information and knowledge management**

Full text available: pdf(974.45 KB) Additional Information: full citation, abstract, references, citings, index terms

In large data- and access-intensive web sites, efficient and reliable access is hard to achieve. This situation gets even worse for web sites providing precise structured query facilities and requiring topicality of the presented information even in face of a highly dynamic content. The achievement of these partly conflicting goals is strongly influenced by the approach chosen for page generation, ranging from composing a web page upon a user's request to its generation in advance. The offi ...

Keywords: WWW, optimization, page generation, reliability, tourism information system

16 The scent of a site: a system for analyzing and predicting information scent, usage, and usability of a Web site
Ed H. Chi, Peter Pirolli, James Pitkow
April 2000 **Proceedings of the SIGCHI conference on Human factors in computing systems**

Full text available: pdf(1.29 MB) Additional Information: full citation, abstract, references, citings, index terms

Designers and researchers of users' interactions with the World Wide Web need tools that permit the rapid exploration of hypotheses about complex interactions of user goals, user behaviors, and Web site designs. We present an architecture and system for the analysis and prediction of user behavior and Web site usability. The system integrates research on human information foraging theory, a reference model of information visualization and Web

data-mining techniques. The system also incorporat ...

Keywords: World Wide Web, data mining, dome tree, information foraging, information scent, information visualization, longest repeated subsequences, usability, usage-based layout

17 Acrophile: an automated acronym extractor and server

Leah S. Larkey, Paul Ogilvie, M. Andrew Price, Brenden Tamilio

June 2000 **Proceedings of the fifth ACM conference on Digital libraries**

Full text available: pdf(118.52 KB) Additional Information: full citation, abstract, references, citings, index terms

We implemented a web server for acronym and abbreviation lookup, containing a collection of acronyms and their expansions gathered from a large number of web pages by a heuristic extraction process. Several different extraction algorithms were evaluated and compared. The corpus resulting from the best algorithm is comparable to a high-quality hand-crafted site, but has the potential to be much more inclusive as data from more web pages are processed.

Keywords: acronyms, information extraction

18 Finding and visualizing inter-site clan graphs

Loren Terveen, Will Hill

January 1998 **Proceedings of the SIGCHI conference on Human factors in computing systems**

Full text available: pdf(1.16 MB) Additional Information: full citation, references, citings, index terms

Keywords: co-citation analysis, collaborative filtering, computer supported cooperative work, human-computer interaction, information access, information retrieval, information visualization, social filtering, social network analysis

19 Supporting classroom information management with SCOUT

Quranna Khan, D. Scott McCrickard, Sherian Clay

April 1999 **Proceedings of the 37th annual Southeast regional conference (CD-ROM)**

Full text available: pdf(44.71 KB) Additional Information: full citation, index terms

20 Beyond document similarity: understanding value-based search and browsing technologies

Andreas Paepcke, Hector Garcia-Molina, Gerard Rodriguez-Mula, Junghoo Cho

March 2000 **ACM SIGMOD Record**, Volume 29 Issue 1

Full text available: pdf(1.29 MB) Additional Information: full citation, abstract, citings, index terms

In the face of small, one or two word queries, high volumes of diverse documents on the Web are overwhelming search and ranking technologies that are based on document similarity measures. The increase of multimedia data within documents sharply exacerbates the shortcomings of these approaches. Recently, research prototypes and commercial experiments have added techniques that augment similarity-based search and ranking. These techniques rely on judgments about the 'value' of documents. Jud ...

Keywords: World-Wide Web, collaborative filtering, hypertext, information filters, information retrieval, links, metadata, ranking, relevance, search engines

Results 1 - 20 of 52 Result page: **1** 2 3 next

Published before December 2000
Terms used **dynamic page crawl hyperlink**

Found **36** of **111,743**

Sort results by: relevance
Display results: expanded form
Save results to a Binder
Search Tips
Open results in a new window

Try an Advanced Search
Try this search in The ACM Guide

Results 1 - 20 of 36 Result page: **1** 2 next

Relevance scale

1 The scent of a site: a system for analyzing and predicting information scent, usage, and usability of a Web site
Ed H. Chi, Peter Pirolli, James Pitkow
April 2000 **Proceedings of the SIGCHI conference on Human factors in computing systems**

Full text available: pdf(1.29 MB) Additional Information: full citation, abstract, references, citings, index terms

Designers and researchers of users' interactions with the World Wide Web need tools that permit the rapid exploration of hypotheses about complex interactions of user goals, user behaviors, and Web site designs. We present an architecture and system for the analysis and prediction of user behavior and Web site usability. The system integrates research on human information foraging theory, a reference model of information visualization and Web data-mining techniques. The system also incorporat ...

Keywords: World Wide Web, data mining, dome tree, information foraging, information scent, information visualization, longest repeated subsequences, usability, usage-based layout

2 Hypertext data mining (tutorial AM-1)
Soumen Chakrabarti
August 2000 **Tutorial notes of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available: pdf(1.08 MB) Additional Information: full citation, index terms

3 Finding replicated Web collections
Junghoo Cho, Narayanan Shivakumar, Hector Garcia-Molina
May 2000 **ACM SIGMOD Record , Proceedings of the 2000 ACM SIGMOD international conference on Management of data**, Volume 29 Issue 2

Full text available: pdf(332.72 KB) Additional Information: full citation, abstract, references, citings, index terms

Many web documents (such as JAVA FAQs) are being replicated on the Internet. Often entire document collections (such as hyperlinked Linux manuals) are being replicated many times. In this paper, we make the case for identifying replicated documents and collections

to improve web crawlers, archivers, and ranking functions used in search engines. The paper describes how to efficiently identify replicated documents and hyperlinked document collections. The challenge is to identify these replicas ...

4 Crawler-Friendly Web Servers
Onn Brandman, Junghoo Cho, Hector Garcia-Molina, Narayanan Shivakumar
September 2000 **ACM SIGMETRICS Performance Evaluation Review**, Volume 28 Issue 2
Full text available: pdf(513.04 KB) Additional Information: full citation, abstract, index terms

In this paper we study how to make web servers (e.g., Apache) more crawler friendly. Current web servers offer the same interface to crawlers and regular web surfers, even though crawlers and surfers have very different performance requirements. We evaluate simple and easy-to-incorporate modifications to web servers so that there are significant bandwidth savings. Specifically, we propose that web servers export meta-data archives decribing their content.

5 Topical locality in the Web
Brian D. Davison
July 2000 **Proceedings of the 23rd annual international ACM SIGIR conference on Research and development in information retrieval**
Full text available: pdf(771.77 KB) Additional Information: full citation, abstract, references, citings, index terms

Most web pages are linked to others with related content. This idea, combined with another that says that *text in, and possibly around, HTML anchors describe the pages to which they point*, is the foundation for a usable World-Wide Web. In this paper, we examine to what extent these ideas hold by empirically testing whether topical locality mirrors spatial locality of pages on the Web. In particular, we find that the likelihood of linked pages having similar textual content to be ...

6 Organizing topic-specific web information
Sougata Mukherjea
May 2000 **Proceedings of the eleventh ACM on Hypertext and hypermedia**
Full text available: pdf(183.02 KB) Additional Information: full citation, references, citings, index terms

Keywords: World-Wide Web, abstraction hierarchy, graph algorithms, information visualization, topic management

7 Integrating content search with structure analysis for hypermedia retrieval and management
Wen-Syan Li, K. Selçuk Candan
December 1999 **ACM Computing Surveys (CSUR)**
Full text available: pdf(25.42 KB) Additional Information: full citation, references, index terms

Keywords: link analysis, organization, topic distillation

8 Information retrieval on the web
Mei Kobayashi, Koichi Takeda
June 2000 **ACM Computing Surveys (CSUR)**, Volume 32 Issue 2

Full text available: pdf(213.89 KB) Additional Information: full citation, abstract, references, citings, index terms

In this paper we review studies of the growth of the Internet and technologies that are useful for information search and retrieval on the Web. We present data on the Internet from several different sources, e.g., current as well as projected number of users, hosts, and Web sites. Although numerical figures vary, overall trends cited by the sources are consistent and point to exponential growth in the past and in the coming decade. Hence it is not surprising that about 85% of Internet user ...

Keywords: Internet, World Wide Web, clustering, indexing, information retrieval, knowledge management, search engine

9 Data mining and the Web: past, present and future

Minos N. Garofalakis, Rajeev Rastogi, S. Seshadri, Kyuseok Shim

November 1999 **Proceedings of the 2nd international workshop on Web information and data management**

Full text available: pdf(660.55 KB) Additional Information: full citation, references, citings, index terms

10 Indexing and retrieval of scientific literature

Steve Lawrence, Kurt Bollacker, C. Lee Giles

November 1999 **Proceedings of the eighth international conference on Information and knowledge management**

Full text available: pdf(985.22 KB) Additional Information: full citation, abstract, references, citings, index terms

The web has greatly improved access to scientific literature. However, scientific articles on the web are largely disorganized, with research articles being spread across archive sites, institution sites, journal sites, and researcher homepages. No index covers all of the available literature, and the major web search engines typically do not index the content of Postscript/PDF documents at all. This paper discusses the creation of digital libraries of scientific literature on the web, incl ...

11 Constructing, organizing, and visualizing collections of topically related Web resources

Loren Terveen, Will Hill, Brian Amento

March 1999 **ACM Transactions on Computer-Human Interaction (TOCHI)**, Volume 6 Issue 1

Full text available: pdf(303.62 KB) Additional Information: full citation, abstract, references, citings, index terms

For many purposes, the Web page is too small a unit of interaction and analysis. Web sites are structured multimedia documents consisting of many pages, and users often are interested in obtaining and evaluating entire collections of topically related sites. Once such a collection is obtained, users face the challenge of exploring, comprehending and organizing the items. We report four innovations that address these user needs: (1) we replaced the Web page with the Web site

Keywords: cocitation analysis, collaborative filtering, computer supported cooperative work, information visualization, social filtering, social network analysis

12 Defining logical domains in a web site

Wen-Syan Li, Okan Kolak, Quoc Vu, Hajime Takano

May 2000 **Proceedings of the eleventh ACM on Hypertext and hypermedia**

Full text available: pdf(152.26 KB) Additional Information: full citation, references, citings, index terms

Keywords: WWW, domain boundary, link structures, logical domain, site map

13 Of crawlers, portals, mice, and men: is there more to mining the Web?
Minos N. Garofalakis, Sridhar Ramaswamy, Rajeev Rastogi, Kyuseok Shim
June 1999 **ACM SIGMOD Record , Proceedings of the 1999 ACM SIGMOD international conference on Management of data**, Volume 28 Issue 2
Full text available: pdf(148.19 KB) Additional Information: full citation, abstract, index terms

The World Wide Web is rapidly emerging as an important medium for transacting commerce as well as for the dissemination of information related to a wide range of topics (e.g., business, government, recreation). According to most predictions, the majority of human information will be available on the Web in ten years. These huge amounts of data raise a grand challenge for the database community, namely, how to turn the Web into a more useful information utility. This is exactly the subject t ...

14 Tools and approaches for developing data-intensive Web applications: a survey
Piero Fraternali
September 1999 **ACM Computing Surveys (CSUR)**, Volume 31 Issue 3
Full text available: pdf(524.80 KB) Additional Information: full citation, abstract, references, citings, index terms

The exponential growth and capillar diffusion of the Web are nurturing a novel generation of applications, characterized by a direct business-to-customer relationship. The development of such applications is a hybrid between traditional IS development and Hypermedia authoring, and challenges the existing tools and approaches for software production. This paper investigates the current situation of Web development tools, both in the commercial and research fields, by identifying and characte ...

Keywords: HTML, Intranet, WWW, application, development

15 Improved algorithms for topic distillation in a hyperlinked environment
Krishna Bharat, Monika R. Henzinger
August 1998 **Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval**
Full text available: pdf(1.15 MB) Additional Information: full citation, references, citings, index terms

16 Survey articles: Web usage mining: discovery and applications of usage patterns from Web data
Jaideep Srivastava, Robert Cooley, Mukund Deshpande, Pang-Ning Tan
January 2000 **ACM SIGKDD Explorations Newsletter**, Volume 1 Issue 2
Full text available: pdf(1.44 MB) Additional Information: full citation, abstract, references, citings

Web usage mining is the application of data mining techniques to discover usage patterns from Web data, in order to understand and better serve the needs of Web-based applications. Web usage mining consists of three phases, namely *preprocessing, pattern discovery*, and *pattern analysis*. This paper describes each of these phases in detail. Given its application potential, Web usage mining has seen a rapid increase in interest, from both the research and practice communities. This pap ...

Keywords: data mining, web usage mining, world wide web

17 Evolving intelligent text-based agents
Edmund S. Yu, Ping C. Koo, Elizabeth D. Liddy
June 2000 **Proceedings of the fourth international conference on Autonomous agents**
Full text available: pdf(1.14 MB) Additional Information: full citation, references, citings, index terms

Keywords: evolution of agents, information agents, learning and adaptation, multi-agent teams

18 Finding and visualizing inter-site clan graphs
Loren Terveen, Will Hill
January 1998 **Proceedings of the SIGCHI conference on Human factors in computing systems**
Full text available: pdf(1.16 MB) Additional Information: full citation, references, citings, index terms

Keywords: co-citation analysis, collaborative filtering, computer supported cooperative work, human-computer interaction, information access, information retrieval, information visualization, social filtering, social network analysis

19 Mining the Web for acronyms using the duality of patterns and relations
Jeonghee Yi, Neel Sundaresan
November 1999 **Proceedings of the 2nd international workshop on Web information and data management**
Full text available: pdf(564.90 KB) Additional Information: full citation, abstract, references, citings, index terms

The Web is a rich source of information, but this information is scattered and hidden in the diversity of web pages. Search engines are windows to the web. However, the current search engines, designed to identify pages with specified phrases have very limited power. For example, they cannot search for phrases related in a particular way (e.g. books and their authors).In this paper we present a solution for identifying a set of inter-related information on the web using the

20 Database techniques for the World-Wide Web: a survey
Daniela Florescu, Alon Levy, Alberto Mendelzon
September 1998 **ACM SIGMOD Record**, Volume 27 Issue 3
Full text available: pdf(1.79 MB) Additional Information: full citation, citings, index terms

Results 1 - 20 of 36 Result page: **1** 2 next